DOCUMENT RESUME

ED 051 983 RE 003 688

AUTHOR Askov, Warren; And Others
TITLE Assessment of the de Hirsch Predictive Index Tests of Reading Failure.
PUB DATE 22 Apr 71
NOTE 14p.; Paper presented at the meeting of the International Reading Association, Atlantic City, N.J., Apr 19-23, 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Group Tests, Kindergarten Children, Predictive Ability (Testing), *Predictive Validity, *Reading Readiness Tests, Reading Research, *Reading Tests, *Standardized Tests, *Testing Programs, Test Reviews

ABSTRACT

The predictive validity and the general usability of a battery of 10 tests reported by de Hirsch, Jansky, and Langford, the de Hirsch Predictive Index Tests of reading failure, were examined. The de Hirsch battery was administered to 433 kindergarten children in six public schools. When the pupils entered first grade, the Metropolitan Readiness Test was administered; in second grade, pupils took the Gates-MacGinitie Reading Test. Regression analysis was used to assess the degree of additional information the de Hirsch tests added to the prediction of scores on the Gates-MacGinitie, beyond that provided by the Metropolitan. The alphabet subtest of the Metropolitan and the word reversals and word reproduction tests of the de Hirsch contributed the most to the prediction of second-grade comprehension scores. Discriminate analysis showed that the Metropolitan Test used alone properly classified 28 of 29 poor readers and 68 of 79 superior readers, and that the alphabet subtest discriminated nearly as well as all six Metropolitan subtests. It was concluded that it is questionable whether the time and expense involved in administering the de Hirsch tests is justified for general testing, but that they may be useful for assessment of marginal kindergarten students. Tables and references are included. (VJ)

ED051983

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Warren H. Askov

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

ASSESSMENT OF THE DE HIRSCH PREDICTIVE INDEX TESTS OF READING FAILURE

Warren Askov, Wayne Otto and Richard Smith

Research Reports
Thursday, April 22, 4:00-5:00 P.M.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

c 1970 - The Regents of the University of Wisconsin for the Wisconsin Research and Development Center for Cognitive Learning.

Copyright is claimed only during the period of development, test, and evaluation, unless authorization is received from the U.S. Office of Education to claim copyright on the final materials. For the current copyright status contact either the copyright proprietor or the U.S. Office of Education.

RE003 688

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.
Center No. C-03 / Contract OE 5-10-154

Warren Askov
Wisconsin Research and Development
Center for Cognitive Learning
University of Wisconsin
1404 Regent Street
Madison, Wisconsin 53706

ASSESSMENT OF THE DE HIRSCH PREDICTIVE INDEX TESTS OF READING FAILURE

Research Reports
Thursday, April 22, 4:00-5:00 P.M.

OBJECTIVES

The purpose of the present three-year study was to investigate (a) the predictive validity and (b) the general usability in a school setting of a battery of ten tests reported by de Hirsch, Jansky and Langford (1) to be predictors of reading failure.

The need for further study of the de Hirsch predictive battery seemed apparent for a number of reasons. First, the subjects in the de Hirsch study were a select group that did not represent the wide range of mental ability in most kindergarten classes. Second, in the de Hirsch study the predictive battery was administered in a

clinical setting which might produce results unlike those which would be obtained in a more typical school setting. Third, of the 53 subjects in the de Hirsch study only six were judged to be "less than adequate" readers and only five were judged to be "poor" readers at the end of second grade. Furthermore, five of those judged to be "less than adequate" readers at the end of second grade scored at the 2.5 to 3.4 grade level on the Gates Advanced Primary Test and 2.4 or below on the Gray Oral Reading Test, which means that they may have been considered "adequate" readers by other judges' standards. Finally, although de Hirsch, Jansky and Langford report their study to be "preliminary in nature," the battery they have identified has been cited in the professional literature dealing with the prevention and correction of reading disability (2; 3) and is reportedly being used with some modification in public schools. This obvious impatience to report, cite and use the results of a preliminary study attests to the interest in the effort to predict reading failure and emphasizes the need for immediate and thorough investigation of the de Hirsch battery.

METHODS

Subjects

The de Hirsch battery was administered to 433 kindergarten children, average age five years, ten months, at six Madison Public Schools. In order to determine the feasibility of using the de Hirsch tests with children of varied characteristics, no attempt was made to prescreen the children for ability, SES, or other

variables.

Experimental Variables

Kindergarten tests. Data collected in kindergarten consisted of age, sex, scores on the ten tasks of the de Hirsch predictive index, and scores on a word association task which had previously been found to be related to reading ability in older children. Two of the ten tests de Hirsch et al. originally administered individually were group administered in the present study: the Bender Visual-Motor Gestalt test and the Word Matching subtest of the Gates Reading Readiness battery. These two tests were given to intact kindergarten classes by the same examiner and one of two assistants. With the exception of the group administration of these two tests, the content and administration of all tests were essentially similar to the description by de Hirsch et al. The remaining tests, administered individually, involved (1) ability to hold a pencil, (2) Wepman Auditory Discrimination Task, (3) number of words used to tell The Three Bears story, (4) providing generic category names for three groups of words, (5) a word reversals task, and (6) recognition and reproduction of two words previously taught. A word association task was added. In that task, children were asked to give their first association to each of seven words. The associations were scored in terms of consensuality with all associations given. High consensuality received a high score. Each child received a total score for all seven words.

First grade tests. All subjects were given the Metropolitan Readiness Test at the beginning of Grade 1 (Fall, 1968).

Second Grade Tests. The Gates-MacGinitie Reading Test (Primary

B), which yields Vocabulary and Comprehension scores, was given as a part of the regular school testing program in February of Grade 2 (1970). We were able to obtain scores for 285 of the original Ss; scores were not available for the remaining 148 original Ss. Twenty-nine of the latter Ss were known to have been retained in Grade 1; the remainder had moved, changed schools, or failed to take the test.

The comprehension subtest of the Gates-MacGinitie Reading Test was chosen as a measure of second grade reading performance. Ss were assigned to one of three groups on the basis of performance on that test: (1) Those reading one-half year or more below middle of second grade, (2) those reading three months above or below middle of second grade, and (3) those reading one-half year or more above the middle of second grade. A fourth group consisted of those retained in first grade (N = 29).

RESULTS AND DISCUSSION

Regression Analysis

The effectiveness of the de Hirsch Predictive Index tests for predicting later reading difficulties was assessed in three ways. First, regression analysis was employed to assess the degree of additional information the de Hirsch tests added to the prediction of second grade Vocabulary and Comprehension scores on the Gates-MacGinitie Reading Test beyond that provided by the use of an established test, the Metropolitan Readiness Tests. The relationship of age, sex of subject, and the school attended were added to the regression equation first. As shown in Tables 1 and 2, the age of the child was not

significantly related to either second-grade Vocabulary or Comprehension scores, but the sex of the child (girls predictably did better) and school attended were significantly related to both. (The finding that school attended is significantly related to second grade performance is confounded by differences between the children, differences among the schools, and the fact that different examiners collected data from each of the schools.) Next, scores on each of the Metropolitan Readiness Tests, the total score on the word association test, and each of the de Hirsch Predictive Index tests were added to the regression equation in that order, and the amount of variance and the additional prediction added by each of these scores were analyzed.

For the mathematically inclined who like to add up degrees of freedom, it should be noted that seventeen children were missing one or more of the de Hirsch or Metropolitan test scores and were eliminated from the regression and the discriminate analyses discussed later. Also, only 141 of the 285 children's Three Bears Stories were scored. The Three Bears stories were extremely time consuming to to score, and we decided to analyze a portion of the stories before scoring the rest. The partial F ratios testing the relationship between the number of words used in the Three Bears Stories and Grade 2 Vocabulary and Comprehension scores are too small to allow rejection of the hypothesis that the Three Bears Stories are unrelated to Vocabulary and Comprehension scores ($df = 1,102$; $F = .81$ and $.02$: $p < .37$ and $.88$ respectively). Since this variable seemed to be contributing so little prediction, it was dropped from the analysis

to avoid contending with missing data. The Metropolitan tests added a highly significant degree of prediction of both second-grade Vocabulary and Comprehension scores ($F = 18.04$ and 12.04; $p < .0001$). The word association task added no significant prediction of either score to that allowed by the above variables. Addition of the ten de Hirsch test scores to the regression equation added a slight, but significant increase to the prediction of second-grade Vocabulary ($F = 2.15$; $p < .025$) and Comprehension scores ($F = 1.97$; $p < .05$).

Analysis of the partial F ratios shown in Tables 1 and 2 reveals the differing effectiveness of the individual subtests of the Metropolitan and the de Hirsch tests for predicting second grade reading performance. The alphabet (letter naming) subtest of the Metropolitan and the Word Reversals and Word Reproduction tests of the de Hirsch battery contribute the most to the prediction of second-grade comprehension scores. Nearly significant contributions to prediction of Comprehension scores are added by the Word Meaning and Listening subtests of the Metropolitan and by the Pencil Use and Auditory Discrimination tests of the de Hirsch battery. Individual tests variables significantly related to second grade Vocabulary scores, as shown by the partial F ratios in Table 2, are the Alphabet subtests of the Metropolitan and the pencil use, word reversal, and word reproduction tests of the de Hirsch battery.

Discriminant Analysis

A second statistical strategy employed to compare the predictive effectiveness of the de Hirsch and the Metropolitan Readiness Tests was a test of discriminate analysis. Discriminate analysis allows

the assignment of subjects to one of several groups on the basis of statistical similarity to that group. In this case, two groups were employed, the 29 children who were retained in first grade and the remaining subjects who might be identified as passing readers. As mentioned earlier, the two groups could be classified on the basis of their second grade Gates-MacGinitie vocabulary and comprehension scores as reading at second grade level or as reading one half year or more above or below grade level. If the kindergarten and first grade tests given earlier were effective in predicting later reading achievement, then they should best discriminate those children retained in first grade and those passed to second grade but reading below grade level (i.e. children with reading difficulities) from those students reading at least one-half year above grade level (superior readers). If the tests cannot discriminate between these two extremes of reading performance, then they cannot be expected to discriminate among students reading more nearly at the average for their grade level. Thus, the attempt was to statistically assign each of the children tested to either a high risk or a superior reading group. The fit to the categories actually observed should be best at the two extremes. Assignment of the below average and average readers to either a retained or above average group reflects their statistical similarity to either one or the other of the two groups and should be moderately successful. Thus, if we force assignment of subjects to one of two groups -- poor readers or superior readers -- we should be most successful at the extremes and moderately successful with the below average and average second grade readers.

Table 3 indicates that had we administered the Metropolitan Readiness tests only, we would have properly classified 28 of the 29 children retained in first grade at the one extreme and 68 of the 79 children who were at the other extreme, i.e. reading one-half year or more above grade level in second grade. We would have missed one of the retained students and misclassified 11 of the above average students. Knowledge of the de Hirsch test scores in addition to the Metropolitan scores decreases the number of misassignments that we would make in all but the group of children retained in first grade.

It should be kept in mind that these assignments are based on kindergarten and first grade readiness tests that were given nearly two years before our criterion second grade reading test. If our interest is in identifying those children who may experience later reading difficulty, then by administering the Metropolitan tests we have decreased the number of high risk children by over half -- from 297 to 137. The administration of the de Hirsch tests decreases the original number by two-thirds, from 297 to 98. The success with which the discriminant analysis has assigned nearly all of the retained subjects as well as some of the above average readers to a high risk group indicates that the tests may be somewhat hard for the average kindergarten child. It does pick up most of the failing readers, but also a number of the passing readers. Perhaps, as de Hirsch et. al.(1) observe ". . . in order to increase the chances of identifying virtually all failing children, it is necessary to throw out a large net, as it were, one which will inevitably pick up some adequate readers" (1:42).

Still, one is left with the question whether the amount of time required to administer the de Hirsch tests, most of which must be administered individually and taking about twenty to thirty minutes per child, is justified by an adequate return on the investment. The clinician's and the teacher's time are valuable commodities. It is reasonable to ask which of the many tests which were given to the children best predict later reading performance.

The same discriminate analysis described above was employed to partially answer that question. The program was requested to pick in serial order from all the dependent variables the ones which best allowed assignment to either a retained or an above average reading group. The best five variables in order of the ability to discriminate later reading ability are the (1) alphabet subtest of the Metropolitan Readiness tests and (2) Word Reversal, (3) Bender Visual-Motor Gestalt, (4) Word Recognition I, and (5) Auditory Discrimination tests from the de Hirsch battery. These tests accounted for most of the variance, and the other tests increased discrimination only slightly. As shown in Table 3, if we had administered only the Alphabet subtest, we would have discriminated nearly as well as with all six subtests of the Metropolitan Readiness Test. If we had administered only the Alphabet, word-reversal, and Bender Visual-Motor Gestalt test, we would have discriminated somewhat better than with all Metropolitan tests and only slightly worse than with the Metropolitan and de Hirsch tests combined.

One additional observation should be made. Although both the de Hirsch and Metropolitan tests allow signigicant prediction of

second-grade reading scores, the standard error measurement associated with each measure is sufficiently high to preclude very effective prediction of individual cases (cf. Tables 1 and 2). The de Hirsch tests decrease that standard error of measurement only slightly.

CONCLUSIONS

The significant prediction of second grade reading performance allowed by the de Hirsch tests given in kindergarten nearly two years before the criterion reading measure points to a developmental consistency that heightens the practical significance for examining these and similar developmental tasks at kindergarten. Thus, the de Hirsch tasks may be useful as tests of developmental skills which should be examined in kindergarten, In addition, they do provide additional prediction of later reading ability beyond that allowed by the Metropolitan tests. It is still questionable, however, whether the time and expense of administering them individually is justified for general testing. They may be very useful, though, in providing additional information for the assessment of marginal kindergarten students, thereby supplementing teacher judgment.

REFERENCES

1. de Hirsch, Katrina, Jeannette J. Jansky, and William S. Langford. Predicting Reading Failure: A Preliminary Study. New York: Harper and Row, 1966.

2. Harris, Albert J. How to Increase Reading Ability (5th ed.). New York: McKay, 1970.

3. Strang, Ruth. Diagnostic Teaching of Reading. (2nd ed.). New York: McGraw-Hill, 1969.

Table 1

Summary of Regression Analysis Using Second Grade Gates-MacGinitie Comprehension Test Scores as a Criterion

Source	df	Total SS	SS Added	F	p	MS_e	R	S.E. $\hat{Y}$
Age	1	1,577.71	1,577.71	2.53	.25	834.79	.08	28.893
Sex	1	4,160.72	2,583.01	4.15	.05	828.19	.14	28.778
School	5	11,300.15	7,139.43	2.29	.05	816.66	.22	28.577
School x Sex	5	16,188.32	4,888.17	1.57	.25	813.50	.27	28.522
Metropolitan Tests	6	61,219.09	45,030.77	12.04	.0001	652.26	.52	25.539
Word Meaning*	1			1.99	.16			
Listening*	1			1.95	.16			
Matching*	1			0.13	.72			
Alphabet*	1			8.49	.004			
Numbers*	1			1.04	.31			
Copying*	1			0.53	.47			
Total Word Association	1	63,066.98	1,847.89	2.97	.10	647.44	.53	25.445
de Hirsch Tests	10	75,330.01	12,263.03	1.97	.05	623.12	.58	24.962
Bender*	1			0.61	.43			
Gates Word Matching*	1			0.00	.96			
Pencil Use*	1			2.64	.11			
Wepman*				2.09	.15			
X Errors*	1			0.01	.94			
Y Errors*	1			0.45	.50			
Categories*	1			3.52	.06			
Word Reversals*	1			0.87	.35			
Word Recognition I*	1			1.50	.22			
Word Recognition II*	1			8.34	.004			
Word Reproduction*	1							
Residual	238					623.12		

* Note: The values tabled for the individual tests of the Metropolitan and de Hirsch battery are partial F ratios -- i.e. the relation of that variable to the criterion after the entire model has been fitted.

Table 2

Summary of Regression Analysis Using Second Grade Gates-MacGinitie Vocabulary Test Scores as a Criterion

Source	df	Total SS	SS Added	F	P	MSe	R	S.E. $\hat{Y}$
Age	1	865.70		1.50	.25	847.60	.06	29.114
Sex	1	3,285.38	2,392.88	4.15	.05	841.76	.12	29.013
School	5	11,283.16	8,024.58	2.78	.05	827.09	.22	28.759
School x Sex	5	14,332.65	3,049.49	1.06	.50	831.35	.25	28.833
Metropolitan Tests	6	76,720.04	62,387.40	18.04	.0001	600.83	.58	24.512
Word Meaning*	1			1.70	.19			
Listening*	1			2.70	.10			
Matching*	1			.01	.93			
Alphabet*	1			11.37	.0009			
Numbers*	1			3.44	.06			
Copying*	1			3.11	.07			
Total Word Association	1	76,739.89	19.85			603.17	.58	24.560
de Hirsch Tests	10	89,145.03	12,405.14			576.39	.63	24.008
Bender*	1			0.51	.48			
Gates Word Matching*	1			0.61	.44			
Pencil Use*	1			6.57	.01			
Wepman*	1							
X Errors*	1			3.23	.07			
Y Errors*	1			0.43	.51			
Categories*	1			1.15	.28			
Word Reversals*	1			4.24	.04			
Word Recognition I*	1			0.20	.65			
Word Recognition II*	1			0.55	.45			
Word Reproduction*	1			4.37	.04			
Residual	238					576.39		

* Note: The values tabled for the individual tests of the Metropolitan and de Hirsch battery are partial F ratios -- i.e. the relation of that variable to the criterion after the entire model has been fitted.

Table 3

The Number of Second-Grade Children at Each of Four Reading Levels Who Were Classified by Discriminate Analysis or as Superior Readers on the Basis of their Metropolitan and de Hirsch Readiness Test Scores

Dependent Variables	Group Assigned to	Second Grade Gates-MacGinitie Comprehension Performance				
		Retained	Below Grade Level	At Grade Level	Above Grade Level	Total
Metropolitan Readiness Tests Only	High Risk	28	43	55	11	137
	Superior	1	35	56	68	160
Metropolitan and de Hirsch Tests	High Risk	27	30	36	5	98
	Superior	2	48	75	74	199
Alphabet subtest only	High Risk	28	43	55	13	139
	Superior	1	35	56	66	158
Alphabet, Word Reversals and Bender tests only	High Risk	27	39	38	10	114
	Superior	2	39	73	69	183